PATENT ABSTRACTS OF JAPAN

(11)Publication number :　　　　07-184160

(43)Date of publication of application : 21.07.1995

(51)Int.Cl.	H04N 5/91
	H04N 5/781

(21)Application number : 05-328033　　　　(71)Applicant : CANON INC

(22)Date of filing :　　　24.12.1993　　　　(72)Inventor : YAMAGAMI MIGAKU

(54) DEVICE FOR PROCESSING PICTURE DATA AND AUDIO DATA

(57)Abstract:

PURPOSE: To easily edit and record the expression in relation among a picture file, an audio file and a text file by designating audio data stored in a recording medium, retrieving the picture data relating to the audio data as file information and reading the data.

CONSTITUTION: When a signal processing CPU 13 detects a recording mode transition instruction of an operation section 15 by the user, the CPU 13 executes the following processing and displays a recorded picture to a video output section 23. A mechanism operation section CPU 4 and a drive circuit 5 control a lens system. The circuit 5 drives a shutter 2 based on a control variable from the CPU 4. Furthermore, a stroboscopic lamp 3 is lighted for the image pickup. A light of an object is made incident on an image pickup element 6 via a processing circuit 9 and an A/D converter section 10. A video signal of the element 6 is converted into a standard component video signal by a processing section 19 via a processing circuit 9 and an A/D converter section 10 and the video signal is inputted to a controller 102. The controller 102 displays a moving picture on a video output section 23 via a buffer memory 12 and a D/A converter 26 under the control of the control CPU 13 and records the picture on a recording medium 101 via an I/F 104.

LEGAL STATUS

[Date of request for examination]   25.12.2000

[Date of sending the examiner's decision of
rejection]

[Kind of final disposal of application other than
the examiner's decision of rejection or
application converted registration]

[Date of final disposal for application]

[Patent number]   3337798

[Date of registration]   09.08.2002

[Number of appeal against examiner's
decision of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application] An image and voice process this invention and it relates to the equipment which performs record playback.
[0002]
[Description of the Prior Art] Conventionally, an image and voice data are processed and the still video format method is learned as specification to record. FM record of this conventional still video format is carried out on the truck where image data and voice data are different respectively. Moreover, the field for recording the track number of a related image truck is established in the voice truck by the control code, and it enables it to record the comment (identification information) about a specific image.
[0003]
[Problem(s) to be Solved by the Invention] However, there are the following problems in the above conventional digital electronic cameras.
** That is, although assignment of an image truck to refer to from a voice truck can be performed by the still video format method, a related voice truck cannot be specified from an image truck. Therefore, since all the trucks on the still video floppy which is a record medium in order to decide the relation of voice data and image data had to be searched, the image truck which wants to reproduce and refer to all voice trucks had to be found out and great time amount was needed inevitably, there was a fault that there was much futility and it was not realistic.
[0004] ** When it was two or more images, one voice was not able to be made to correspond again, since only one image can be referred to from a voice truck. Therefore, there was a fault that it could not do at once to two or more images even if it is going to expound on a common concept with voice.
** Since great storage capacity was further needed in order to record voice data, especially the thing for which a voice truck performs a simple comment was seen in cost, and had the problem of being unsuitable.
[0005] In addition, the trouble about the conventional digital electronic camera mentioned above is also a trouble concerning the equipment which processes general voice and general image data. therefore , the place which this invention be make in view of the technical problem which mentioned above , and make into the object be offer the equipment which process the voice and the image data which enable the link during each file of at least one or more images , voice , and a text , constitute each file from a both sides so that it may search at a high speed and may reproduce , can edit easily the expression which have relevance in mutual [between an image , voice , and a text file] , and can record it .
[0006] Moreover, since voice or an image is convertible for text format, it is offering the equipment which processes the voice and the image data which can give a comment (identification information) to an image by small capacity. Moreover, using it as a keyword when searching or creating a database based on the text given as a comment, is also offering the equipment which processes the voice and the image data which become possible.
[0007] Furthermore, processing in which data are automatically linked between the information stored in

other databases etc. is also offering the equipment which processes the voice and the image data which become possible.
[0008]
[Means for Solving the Problem] In order to solve an above-mentioned technical problem and to attain the object, the equipment which processes the image data and voice data of this invention The record medium for associating and recording image data and voice data, It is characterized by providing an assignment means to specify the voice data stored in said record medium, the control means which searches the image data relevant to said specified voice data as file information, and a read-out means to read said searched image data and to reproduce.
[0009] Moreover, a character recognition means to perform character recognition from said image data preferably is provided further, and it is characterized by performing registration processing so that it may relate with each data of said image and voice and can record and search by making the recognition result by this character recognition means into a text file. Moreover, a speech recognition processing means to recognize said voice data as an alphabetic character preferably is provided further, and it is characterized by performing registration processing so that it may relate with each data of said image and voice and can record and search by making the recognition result by this speech recognition processing means into a text file.
[0010] Moreover, a display means to display said each data preferably stored in said record medium as file information is provided further, and it is characterized by displaying edit of the mutual related information during said each file, and the retrieval result of each of said data based on this related information.
[0011]
[Function] As mentioned above, since the equipment which processes the image data and voice data concerning this invention is constituted, the link during each file of at least one or more images, voice, and a text of it is attained, and since it can be searched and reproduced at a high speed from both sides, it can edit a relevant expression easily and can record each file on mutual [between an image, voice, and a text file].
[0012] Moreover, since voice or an image is convertible for text format, a comment (identification information) can be given to an image by small capacity. Moreover, based on the text given as a comment, it searches or it also becomes possible to use it as a keyword when creating a database. Furthermore, processing in which data are automatically linked between the information stored in other databases etc. also becomes possible.
[0013]
[Example] Hereafter, the suitable example of this invention is explained to a detail with reference to an accompanying drawing. Drawing 1 is the block diagram showing the system configuration of the body of a digital electronic camera of this example. In drawing 1 , a record medium 101 is a memory card, a hard disk, etc. based on for example, PCMCIA specification. The voice input circuit 20 and the voice output section 22 are for example, an audio jack or a loudspeaker. A/D converter 24 is equipment which changes a sound signal into a digital signal from an analog signal, and D/A converter 25 is equipment which changes into an analog the digital sound signal sent from CPU13 for signal-processing control. A switch 21 is a selection circuitry which chooses transmission of the sound signal to the voice output section 22. The memory bus controller 102 transmits the image data between the image pick-up signal-processing section 19, DSP13 for signal-processing control, the buffer memory 12 for image display, and the record-medium I/F circuit 104, voice data, etc.
[0014] D/A converter 26 is equipment which changes the digital image data from the buffer memory 12 for image display into an analog video signal, and the video output section 23 is a graphic display device which displays as an image the video signal changed into the analog. 1 is a taking lens, the drawing combination shutter which 2 extracts and serves both as a function and a shutter function, and 3 control each control unit which performs actuation with mechanical stroboscope, mechanism, and CPU4 for control unit control, and the actuation circuit 5 is a circuit which each part of a mechanism system makes drive.

[0015] An image sensor 6 is CCD which changes the reflected light from a photographic subject into an electrical signal, and generates a timing signal required in order that the timing signal generating circuit 7 may operate an image sensor 6 ("TG" is called hereafter). The image sensor actuation circuit 8 is an actuation circuit amplified on the level which can drive [of an image pick-up signal] the signal from the timing signal generating circuit 7, and the front-end processing circuit 9 is equipped with the nonlinear amplifying circuit performed before the CDS circuit for the output noise rejection generated with an image sensor 6, and A/D conversion. A/D converter 10 changes the data after front-end processing into digital one. CPU13 for signal-processing control is DSP for signal-processing control (DIGTAL SIGNAL PROCESSOR) which controls the signal-processing section, the actuation display 14 is a display showing the display for actuation assistance, or the condition of a camera, and control units 15 are input units, such as a keyboard for controlling a camera from the outside. Record-medium I/F104 is a circuit for record-medium I/F for connecting the digital electronic camera and record medium 101 based on this example. And a file format like for example, MS-DOS can be used for the record file format to the record medium of the camera of this example.

[0016] The buffer memory 12 for image display can be accessed in a pixel unit from DSP13 for signal-processing control, and it can draw the actuation means panel of arbitration, displaying the image to photo. If a trackball etc. is mounted in a control unit 15, the same user I/F as GUI (Graphical User Interface) using the pointing device in the personal computer in recent years can be realized. That is, all actuation of a camera is attained because DSP13 for signal-processing control draws various control panels to the buffer for image display and a user operates it with a pointing device (trackball in this case) to that control panel. Moreover, presentation of various information, such as a current condition of the image photoed to the user by DSP13 for signal-processing control by drawing an image, a text, and various graphic forms in the image display section and a camera and file management information on a record medium, is performed. Suppose below that the user I/F actuation by the control unit 15 and Above GUI is called actuation by the control unit 15 in this example. Moreover, suppose that a click or double click actuation by the pointing device etc. is called starting actuation.

[0017] If DSP13 for <monitor of record image at time of image recording mode> signal-processing control detects the recording-mode shift instruction of a user's control unit 15, DSP13 for signal-processing control will display the image which performs and records the following processings on the video output section 23. Control of a lens system is performed by CPU4 for mechanism control unit control, and the mechanism system actuation circuit 5 according to an intention of a photography person. Under the present circumstances, photography conditions etc. are displayed on a control unit 15, and the situation of a camera is told to a photography person. Furthermore, the brightness of a photographic subject is measured by the non-illustrated photometry circuit, and the data showing whenever [drawing / of the drawing combination shutter 2] of a value or shutter speed are derived in CPU for mechanism control unit control. Based on the control value drawn by CPU4 for mechanism control unit control, it extracts by the mechanism system actuation circuit 5, and the combination shutter 2 is driven. Moreover, a stroboscope 3 will be made to emit light depending on the output of a photometry circuit (un-illustrating), and a photograph will be taken. Thus, it is exposed and incidence of the reflected light of a photographic subject is carried out to an image sensor 6 through a taking lens 1 and the drawing combination shutter 2. Under the present circumstances, while the drawing combination shutter 2 restricts the amount of incident light to an image sensor 6, when the interlace read-out mold CCD is used as an image sensor, it is prepared in order to make it incident light not have an adverse effect during a transfer of an image at a signal charge. An image sensor 6 is operated with the driving signal which made the output from TG7 amplify by the image sensor actuation circuit 8. In addition, TG7 is having the actuation controlled by DSP13 for signal-processing control. The output of the image sensor 6 made to drive as mentioned above is sent to the front-end processing circuit 9. In the front-end processing circuit 9, in order to use effectively the D range (digitized signal data) of A/D converter 10 with the CDS processing which removes the low-pass noise generated from an image sensor to an output, processing which makes an image pick-up output nonlinear is performed. The image pick-up signal output by which front-end processing was carried out is changed into a digital signal in A/D

converter 10, is changed into standard component video (for example, a luminance signal, two color-difference signals, and an RGB code) by the image pick-up signal-processing section 19, and is inputted into the memory controller 102. By the memory controller 102, the image pick-up signal digitized by control of DSP13 for signal-processing control is continuously transmitted to buffer memory 12.

[0018] Drawing 2 is drawing showing the display format in image recording mode. The display image outputted to the image display section 23 is divided and displayed in the monitor area part 202 of a record image, and the control panel part 203 of the parameter which users, such as a condition, photography conditions, etc. of a camera, can set up, as shown in drawing 2 . By A/D converter 26, the data by which the buffer memory 12 for image display was digitized are changed into an analog video signal, are outputted to the video output section 23, and are displayed in the video output section 23. A user can check in the monitor area part 202 by using the image to record as a dynamic image.

[0019] The dynamic image caught by the image sensor 6 is displayed on the video output section 23 by the above processing.

If photography is directed to a camera when a <record of image> photography person controls a control unit 15, after DSP13 for signal-processing control accesses the image data which suspended animation display and was held at the buffer memory for a display through a memory controller and performs digital compression processing, it will be recorded on a record medium 101 through record-medium I/F104.

[0020] The image which will be recorded if the writing to the buffer memory 12 for image display is stopped stands it still, and this transfer period and the memory bus controller 102 in an after [record termination] fixed period are displayed on the video output section 23. Therefore, a user can check the static image recorded just now in the video output section 23. Moreover, in order to check after photography the contents of the image recorded more quickly, DSP13 for signal-processing control can add the infanticide image of the compressed image to a compressed file. For example, even if it adds the image thinned out in all directions [of the original image / about 1/8], file capacity does not increase so much. In addition, this image is called an index image.

[0021] At the time of the monitor of the voice at the time of voice record, and a <record> recording mode, the monitor of DSP13 for signal-processing control can be carried out [voice / which connects and records the output of the voice input circuit 20 on the input of the voice output section 22 in a switching circuit 21]. If DSP13 for signal-processing control detects the voice record instruction by a user's control unit 15, DSP13 for signal-processing control will transmit the data changed into the digital data to record-medium I/F104 through reception and the memory bus controller 102 with A/D converter 24.

[0022] DSP13 for signal-processing control ends audio record the bottom wholly as discharge of voice record at the event by DSP13 for signal-processing control having detected discharge of the voice record instruction by a user's control unit 15, or DSP13 for signal-processing control having carried out fixed time amount progress.

While compressing the image and having transmitted to record-medium I/F104 as the above-mentioned <record of an image> explained when recording a <simultaneous record of image and voice> image, and voice simultaneously, DSP13 for signal-processing control saves temporarily the audio data received from A/D converter 24 at the internal buffer of DSP13 for signal-processing control, and when a transfer of an image is completed, it transmits voice data to record-medium I/F104. The period when the video signal called vertical blanking period of 1.4 ms extent does not exist is a vertical-synchronization period and before and after that among the 1 field period 16.7 mses of NTSC system. When transmitting image data at the usual video rate, the image data by which image pick-up signal processing is not carried out between the time amount (about 15 mses) except this vertical blanking period is transmitted. In order to realize this, data are transmitted at the speed of about 10 MByte/sec.

[0023] When sampling voice data by 22kHz by 8 bits of one sample, the data volume for 16.7 mses is about 370 Byte(s). When transmitting this data by the remaining 1.4 mses, the transfer speed of about 260 KByte/sec is required. If such transfer speed is the memory cards based on PCMCIA specification etc. as a record medium, it is satisfying enough rates.

[0024] As explained above, voice data is transmitted during a vertical blanking period among the scan periods of one screen, and the camera of this example is simultaneously recordable by performing and carrying out Time Division Multiplexing of the image transfer to an image period, carrying out [voice] a monitor to an image simultaneously. At this time, since it can be regarded as the information relevant to an image, based on explanation of <the voice to an image file and the link of a text> which mention an image and voice later, file management of the voice data is carried out so that both may refer to each other mutually.

[0025] Moreover, voice data may be transmitted within the level blanking period of a video signal. For example, when sampling voice on 44.1kHz, a stereo, and 8-bit conditions, the voice data of a total of six Byte(s) or 4Byte(s) is transmitted and recorded within 1 time of a level blanking period (1H period). although it is good to record that voice data is interleaved every 1H period in case the medium which has recording mechanisms, such as HDD, as a record medium is used at this time -- the case of semiconductor memory -- image data and voice data -- 1 -- it is not necessary to make it interleave for every H, and voice data is recorded before the image data of 1 field period (1V period) -- as -- 1 -- interleave record in every V may be carried out.

[0026] With the camera of <record of text data> this example, amplification of the storage capacity for the comment to an image is substantially reduced by recognizing five images and voice and making it text data.
If DSP13 for <record by ** character recognition> signal-processing control detects the character recognition recording-mode shift instruction by a user's actuation display 15, a camera will output the image which performs and carries out character recognition of the same processing as <the record monitor at the time of image recording mode> to the image display section. The image displayed on the image display section 301 at this time divides and displays the control panel 304 of the parameter with which a user can set up the image 302 which it is going to recognize like drawing 3 , the part 303 which displays the result by which character recognition was carried out, a condition, photography conditions of a camera, etc.

[0027] And the following processings are repeated while the user is directing character recognition activation by the control unit 15. DSP13 for signal-processing control once stops the writing to the buffer memory for image display, and performs character recognition processing to the image. After the character recognition processing to the image on buffer memory is completed, DSP13 for signal-processing control is displayed on the viewing area 303 which shows the recognition result to drawing 3 .

[0028] A user ends directions of character recognition activation by the control unit 15 in the place where the satisfactory recognition result was obtained. At this event, a user directs decision of the text recognized by the control unit 15. A camera records the fixed text on a record medium 101 through the memory controller 102 and record-medium I/F104. When text data is not decided, the rewriting-in above-mentioned actuation is again repeated for a recognition image to buffer memory.

[0029] If DSP13 for <record by ** speech recognition> signal-processing control detects the character recognition recording-mode shift instruction by a user's control unit 15, a camera will perform the same processing as having explained to <the monitor of the voice at the time of voice record, and record>, and will carry out [voice] a monitor. At this time, the display image of the image display section 401 divides and displays the condition of the part 402 which displays the result by which speech recognition was carried out like drawing 4 , and a camera, the control panel 403 of the parameter which users, such as recognition conditions, can set up, etc. As a still more suitable example, it can constitute like the panel 130 of drawing 13 .

[0030] And while the user is directing speech recognition activation by the control unit 15, DSP13 for signal-processing control performs reception and speech recognition processing for the data changed into the digital data with A/D converter 24, and displays a recognition result on the display 402 of drawing 4 . A user ends directions of speech recognition activation by the control unit 15 in the place where the audio recognition result was obtained.

[0031] When the dissatisfaction is in a recognition result, speech recognition activation is again directed

by the control unit 15. The recognition result to satisfy is obtained, and when it is cod roe, a user directs decision of the text recognized by the control unit 15. A camera records the fixed text on a record medium through the memory controller 102 and record-medium I/F104.

[0032] By the above explanation, the digital electronic camera of this example can hold three kinds of files, image data, voice data, and text data, to a record medium so that clearly.

<List display of files> drawing 5 is drawing showing the display format which displays the list of the recorded files. When indicating three sorts of files by the directory as an image at the display of a camera, DSP13 for signal-processing control reads the file currently recorded on the record medium, and the index image 501 which expresses an image file about an image as shown in drawing 5 according to the class of file data, voice, and a text are displayed as an icon respectively like 502 and 504. These displays can be displayed in order of chart lasting time, and, only as for an image, only voice can also display only a text. Moreover, if a user clicks the elimination carbon button 509 after choosing the icon showing an index image, voice, and a text with a pointing device, a camera will eliminate the selected file.

[0033] It is to show whether the voice in the lower part of the index image 501, the text in which each icon 511 and 512 of a text was linked to the image file, and voice exist. A text and voice are linked according to <the voice to an image file and the link of a text> which are mentioned later. For example, it expresses that data exist by making an icon into a gray level like 511 and 512. These icons are chosen and started using a pointing device, and <audio playback> or <playback of a text> mentioned later is performed.

[0034] In drawing 5 , some screens 510 are the scroll bars for scrolling up and down so that they can be searched, when objects to display, such as an index image and an icon, have not gone into a screen. This functions as the tool for the window display used by a personal computer and a workstation in recent years similarly.

Amplification of an image and <playback> drawing 7 are drawings showing the display format of the expanded image. A user chooses one index image using the pointing device of a control unit and performs an amplification instruction (for example, double click of a carbon button) to carry out amplification playback about one image in drawing 7 . DSP13 for signal-processing control reads compression image data from the image file chosen when the above-mentioned actuation was detected, develops, is transmitted to the buffer memory for image display, and is displayed on the image display section. As for the display at this time, 705, 706, etc. are displayed, respectively as the display 701 of the image expanded as shown in drawing 7 , each carbon buttons 702-704 for control and two or more voice linked further, and an icon of a text.

[0035] If a carbon button 702 is started with a pointing device, the panel of the image displayed on drawing 7 will be closed, and will return to the display of above-mentioned drawing 5 .

In <audio playback>, next drawing 5 , a user chooses and starts a voice icon using the pointing device of a control unit to reproduce voice.

[0036] DSP13 for signal-processing control controls a switching circuit 21, and connects the output from D/A converter 25 to the input of the voice output section 22. DSP13 for signal-processing control controls record-medium I/F104 and the memory bus controller 102, reads voice data, and outputs it to a D/A converter by the sampling period at the time of record, and voice is outputted from the voice output section 22.

[0037] A user chooses and starts a text icon using the pointing device of a control unit to display <playback of a text>, and also a text. Drawing 7 is drawing showing the display format of text data. DSP13 for signal-processing control transmits text data as controlled record-medium I/F104 and the memory bus controller 102, read text data, and carried out character bit pattern expansion, for example, shown in drawing 6 to the buffer memory for image display, and displays it on the image display section.

[0038] In drawing 6 , a display 601, the carbon button 602 for control, etc. of a text are displayed. If a carbon button 602 is started with a pointing device, the panel of drawing 6 will be closed.

The camera based on <link of voice [to an image file] and text> this example is equipped with the

icons 703 and 704 which show the control carbon button for adding the text by voice or speech recognition to an image to drawing 7 , respectively, when one image is reproduced according to explanation by <above-mentioned amplification and playback> of an independent image. If the icons 703 and 704 in drawing 7 are clicked on them and started with a pointing device, DSP13 for signal-processing control will perform voice and text capture like <the monitor of the voice at the time of voice record and record>, and <record by ** speech recognition>, respectively. This actuation can be performed by expanding and reproducing and carrying out the monitor of the image data. Drawing 13 is drawing showing signs that the panel 130 for control of speech recognition was displayed in piles on the image display panel of drawing 7 . For example, the panel 130 for control of speech recognition can be displayed in piles on the image display panel of drawing 7 like drawing 13 . When carrying out simultaneous record of image data and the voice data, and when recording voice and a text, carrying out the monitor of the result of having reproduced image data as mentioned above, it becomes the file linked mutually by storing in a file the data with which both an image file and voice, and a text file refer to each other. Drawing 8 shows the data configuration of the image file in a file, and each voice file and text file.

[0039] In drawing 8 , each file stores the identifier of the file to which the number of files (it is [number / of voice files] k about m and the number of image files in n and the number of text files, respectively) which is linked, and which was linked for every file format and the number of files responded. Since not only a voice file but the image file will refer to voice and a text file if it does in this way, even if it does not search all files like before, the voice and the text file which are related from one image file can be specified, and the file can be reproduced or displayed.

[0040] When the voice and the text which were linked to one image file exist, and displaying the image, the voice file and text file which were linked as shown in drawing 5 and drawing 7 , respectively can be displayed by the icon (in for example, lower part of image data). If a user chooses and starts this icon with a pointing device, a camera will perform playback of the associated voice, and the display of a text.

[0041] A text is displayed like drawing 6 . This display is displayed in piles, after that independence or drawing 5 , and drawing 7 display. Moreover, if a user clicks the deletion carbon button 707 after choosing icons 705 and 706 with a pointing device in drawing 7 , a camera can perform deletion of the associated voice and a text file. This actuation means eliminates each other link information of the image file shown in drawing 8 , voice, and a text file. You may make it the voice and the text file from which the link was deleted at this time exist independently, and may make it eliminate them.

[0042] Moreover, in order to make it display independently, in drawing 5 , it is expressed like the file list display of icons 502 and 504, respectively.
In <a list display of a file> of the <link of grouping [of two or more image files] and voice [to a group], and text> above-mentioned, if two or more images are chosen with a pointing device, a frame will be thickly displayed to emphasize that the index image was chosen like drawing 9 . Furthermore, if the grouping carbon button 901 of drawing 9 is clicked with a pointing device, the group of an image file will be created. In order to emphasize that grouping was carried out at this time, the color of the frame of an index image can also be made into a different color from other images.

[0043] Furthermore, if the voice addition carbon button 902 or the voice text addition carbon button 903 shown in drawing 9 with a pointing device is started, DSP13 for signal-processing control will perform voice and text capture like explanation by <the monitor of the voice at the time of voice record and record>, and <record by ** speech recognition>, respectively. Two or more voice in this event and a text will be added to a group. The comment explanation about a matter peculiar to a group can be given by this actuation.

[0044] Two or more images, and voice and a text are linked by the actuation explained above. If the reference data to all the other files about each file in a group are added at this time, in order to specify this group behind, all files must be searched and management will become difficult. So, in this example, when a group's generating becomes clear, the group file for holding group information is created. The configuration of this file comes to be shown in drawing 10 . That is, the image belonging to a group,

voice, the number of the file of each text, and the identifier of each file are stored. Moreover, the image belonging to a group, voice, and a text file are composed like drawing 11 so that reference to the group to whom each file belongs can be performed. That is, reference to the file which belongs to the group only by each file referring to a group is not performed.

[0045] By such a group's file organization, the file which belongs from a group, an image, voice, a text, and any file to group information, i.e., a group, can be identified at a high speed. For example, when searching for other files which belong to the group from one image file, if the group file which once belongs is gained, the direct acquisition of the identifier of those files can be carried out.

[0046] An image file can be belonged to two or more groups so that clearly from the example explained above. If a group file is generated by the above-mentioned procedure, a camera will display the icon 505 which expresses a group file like drawing 5 in <the chart of a file>. If this icon is double-clicked with a pointing device, the file which belongs to a group like drawing 12 will be displayed. In drawing 12 , if a voice icon and a text icon are clicked on them and started with a pointing device, <audio playback> as mentioned above, respectively, and <the display means of a text> are performed, and it can express as the comment peculiar to this group about a matter, and explanation.

[0047] In order to delete the file belonging to a group from a group, after choosing the file deleted in drawing 12 with a pointing device, deletion is started with a carbon button 121. It is made for the voice and the text file which were deleted from the group to exist independently. When making it make it exist independently, in drawing 5 , it comes to appear in a list display like icons 501, 502, and 504.

[0048] Moreover, the file chosen by setting up the elimination carbon button 122 and starting this carbon button may be eliminated. When deleted or eliminated from a group, of course, the link information between the group file and image which are shown in drawing 10 and drawing 11 , and a voice text data file is eliminated.

[0049] Moreover, in order to add a file to the existing group, the group icon of the file added in drawing 5 and the point to add is chosen with a pointing device, and grouping is started. In these actuation, files other than a group can make a multiple selection. An image file, a voice file, and a text file exist in a record medium independently. It is also possible to carry out grouping of these some. In <the chart of a file>, multiple files are chosen with a pointing device. The group file to which an image, voice, and a text belong by furthermore starting the group or carbon button 508 of drawing 5 with a pointing device is created.

[0050] (Effectiveness of this example) Since a single image or two or more single images, and the link between voice and a text are attained and can reproduce the linked relation at a high speed in this example as explained above, the mutual interpolation-expression of an image, voice, and a text can be edited easily.

[0051] Moreover, since voice or an image is changed into a text, a comment can be attached to an image by small capacity. Moreover, it searches based on the text attached as a comment, or considering as the keyword when forming a database also becomes possible. Processing of linking automatically with the information on the database of further others also becomes possible.

[0052] In addition, this invention can apply the above-mentioned example to what corrected or deformed in the range which does not deviate from the meaning. For example, even if it applies to the system which consists of two or more devices, you may apply to the equipment which consists of one device. Moreover, it cannot be overemphasized that it can apply also when attained by supplying a program to a system or equipment.

[0053]
[Effect of the Invention] Since according to this invention the link during each file of at least one or more images, voice, and a text is attained and each file can be searched and reproduced at a high speed from both sides as explained above, there is effectiveness which can edit easily the expression which has relevance in mutual [between an image, voice, and a text file], and can record it.

[0054] Moreover, since voice or an image is convertible for text format, there is effectiveness which can give a comment (identification information) to an image by small capacity. Moreover, based on the text given as a comment, it searches or is effective also in it becoming possible to use it as a keyword when

creating a database. Furthermore, it is effective in processing in which data are automatically linked between the information stored in other databases etc. becoming possible.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any
damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

TECHNICAL FIELD

[Industrial Application] An image and voice process this invention and it relates to the equipment which performs record playback.

[Translation done.]

* NOTICES *

**JPO and INPIT are not responsible for any
damages caused by the use of this translation.**

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

PRIOR ART

[Description of the Prior Art] Conventionally, an image and voice data are processed and the still video format method is learned as specification to record. FM record of this conventional still video format is carried out on the truck where image data and voice data are different respectively. Moreover, the field for recording the track number of a related image truck is established in the voice truck by the control code, and it enables it to record the comment (identification information) about a specific image.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any
damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

EFFECT OF THE INVENTION

(Effectiveness of this example) Since a single image or two or more single images, and the link between voice and a text are attained and can reproduce the linked relation at a high speed in this example as explained above, the mutual interpolation-expression of an image, voice, and a text can be edited easily. [0051] Moreover, since voice or an image is changed into a text, a comment can be attached to an image by small capacity. Moreover, it searches based on the text attached as a comment, or considering as the keyword when forming a database also becomes possible. Processing of linking automatically with the information on the database of further others also becomes possible.
[0052] In addition, this invention can apply the above-mentioned example to what corrected or deformed in the range which does not deviate from the meaning. For example, even if it applies to the system which consists of two or more devices, you may apply to the equipment which consists of one device. Moreover, it cannot be overemphasized that it can apply also when attained by supplying a program to a system or equipment.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention] However, there are the following problems in the above conventional digital electronic cameras.
** That is, although assignment of an image truck to refer to from a voice truck can be performed by the still video format method, a related voice truck cannot be specified from an image truck. Therefore, since all the trucks on the still video floppy which is a record medium in order to decide the relation of voice data and image data had to be searched, the image truck which wants to reproduce and refer to all voice trucks had to be found out and great time amount was needed inevitably, there was a fault that there was much futility and it was not realistic.
[0004] ** When it was two or more images, one voice was not able to be made to correspond again, since only one image can be referred to from a voice truck. Therefore, there was a fault that it could not do at once to two or more images even if it is going to expound on a common concept with voice.
** Since great storage capacity was further needed in order to record voice data, especially the thing for which a voice truck performs a simple comment was seen in cost, and had the problem of being unsuitable.
[0005] In addition, the trouble about the conventional digital electronic camera mentioned above is also a trouble concerning the equipment which processes general voice and general image data. therefore , the place which this invention be make in view of the technical problem which mentioned above , and make into the object be offer the equipment which process the voice and the image data which enable the link during each file of at least one or more images , voice , and a text , constitute each file from a both sides so that it may search at a high speed and may reproduce , can edit easily the expression which have relevance in mutual [between an image , voice , and a text file] , and can record it .
[0006] Moreover, since voice or an image is convertible for text format, it is offering the equipment which processes the voice and the image data which can give a comment (identification information) to an image by small capacity. Moreover, using it as a keyword when searching or creating a database based on the text given as a comment, is also offering the equipment which processes the voice and the image data which become possible.
[0007] Furthermore, processing in which data are automatically linked between the information stored in other databases etc. is also offering the equipment which processes the voice and the image data which become possible.

[Translation done.]

* NOTICES *

OPERATION

[Function] As mentioned above, since the equipment which processes the image data and voice data concerning this invention is constituted, the link during each file of at least one or more images, voice, and a text of it is attained, and since it can be searched and reproduced at a high speed from both sides, it can edit a relevant expression easily and can record each file on mutual [between an image, voice, and a text file].
[0012] Moreover, since voice or an image is convertible for text format, a comment (identification information) can be given to an image by small capacity. Moreover, based on the text given as a comment, it searches or it also becomes possible to use it as a keyword when creating a database. Furthermore, processing in which data are automatically linked between the information stored in other databases etc. also becomes possible.
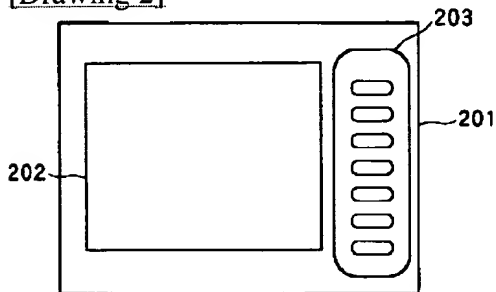

[Translation done.]

* NOTICES *

**JPO and INPIT are not responsible for any
damages caused by the use of this translation.**

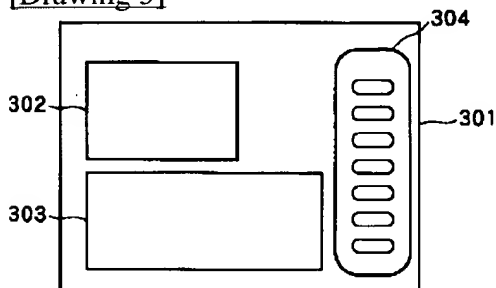1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
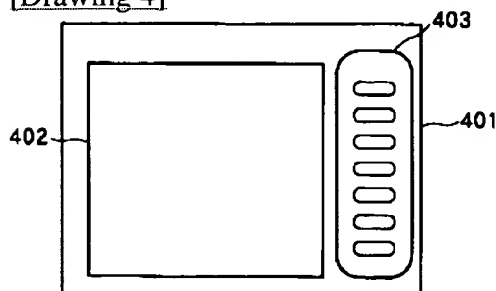3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1] It is the block diagram showing the whole digital electronic camera configuration of the example of this invention.
[Drawing 2] It is drawing showing the display format in image recording mode.
[Drawing 3] It is drawing showing the display format in character recognition mode.
[Drawing 4] It is drawing showing the display format in speech recognition mode.
[Drawing 5] It is drawing showing the list display format of the recorded file.
[Drawing 6] It is drawing showing the display format of a text.
[Drawing 7] It is drawing showing the display format of an image.
[Drawing 8] It is drawing showing the organization inside the file of both sides when an image file, and a voice file and a text file are linked mutually.
[Drawing 9] It is drawing showing a display format when multiple files are chosen in a list display.
[Drawing 10] It is the internal data organization of a group file.
[Drawing 11] They are the image belonging to a group, voice, and drawing showing the data organization inside text each file.
[Drawing 12] It is drawing showing the list display format of the file belonging to a group.
[Drawing 13] It is drawing showing the display format when displaying a speech recognition control panel in piles on an image display format.
[Description of Notations]
1 Taking Lens
6 Image Sensor
10 24 A/D converter
12 Buffer Memory
13 DSP for Signal-Processing Control
14 Actuation Display
15 Control Unit
19 Image Pick-up Signal-Processing Section
20 Voice Input Circuit
22 Voice Output Section
23 Video Output Section
25 26 D/A converter
101 Record Medium
102 Memory Bus Controller
104 Record-Medium I/F

[Translation done.]

* NOTICES *

DRAWINGS

[Drawing 2]

[Drawing 3]

[Drawing 4]
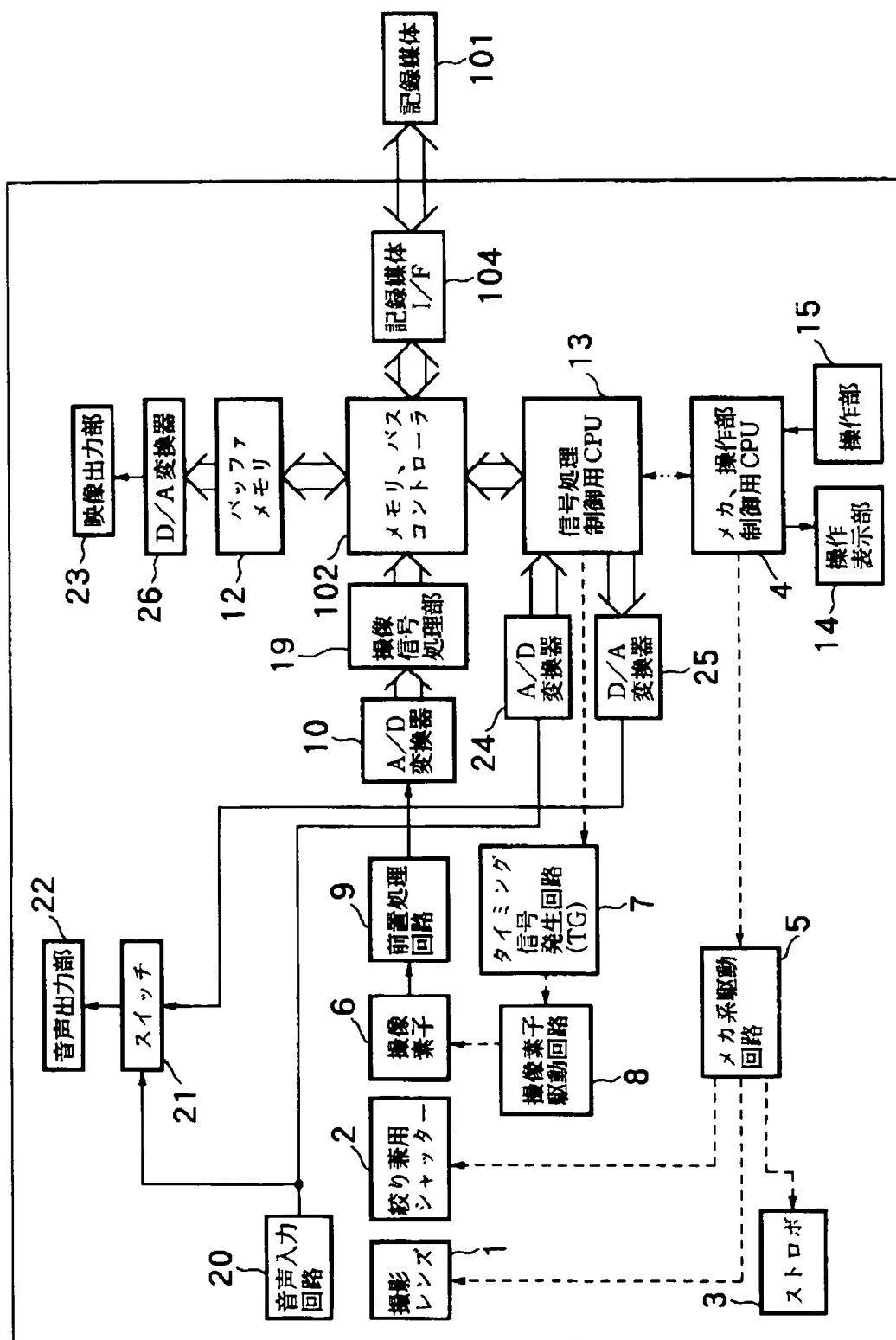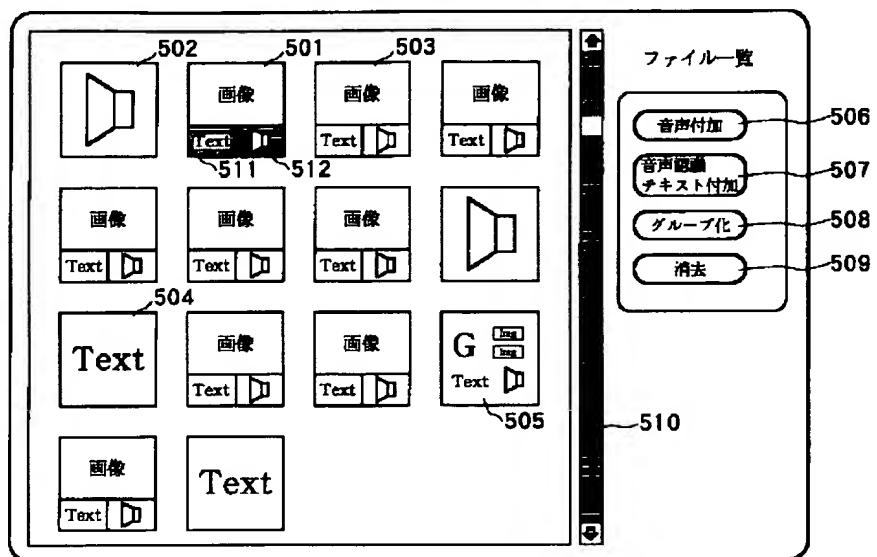
[Drawing 11]

画像ファイル、
音声ファイル、
テキストファイル

リンクされたグループファイルの個数 p
グループファイル1の識別子
グループファイル2の識別子
グループファイル p の識別子

[Drawing 1]

記録媒体 101
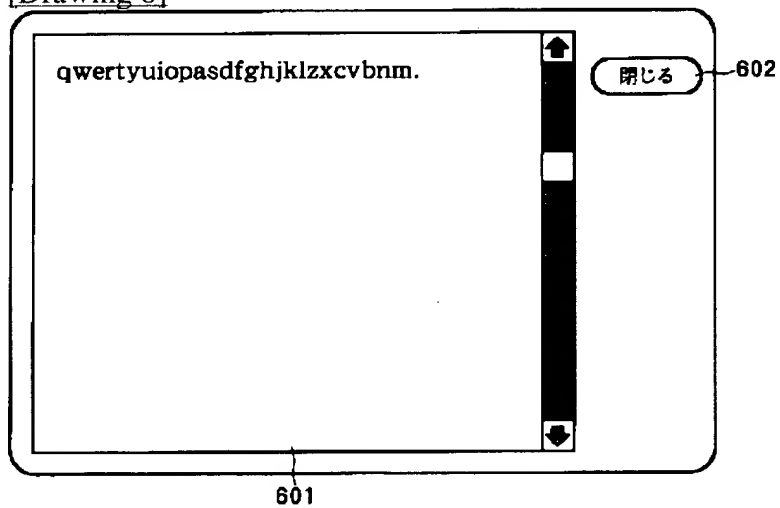
記録媒体I/F 104

映像出力部 23
D/A変換器 26
バッファメモリ 12
メモリ、バスコントローラ 102
撮像信号処理部 19
A/D変換器 10
前置処理回路 9
撮像素子 6
絞り兼用シャッター 2
撮影レンズ 1

信号処理制御用CPU 13
A/D変換器 24
D/A変換器 25
タイミング信号発生回路(TG) 7
撮像素子駆動回路 8

メカ、操作部制御用CPU 4
操作部 15
操作表示部 14

メカ系駆動回路 5
ストロボ 3

音声出力部 22
スイッチ 21
音声入力回路 20

[Drawing 5]

502 501 503

画像 画像 画像

Text 511 512

音声付加 —506

音声図語
テキスト付加 —507

画像 画像 画像

Text Text Text

グループ化 —508

消去 —509

504

Text 画像 画像

Text Text

G 画像
画像
Text
505

510

画像

Text

Text
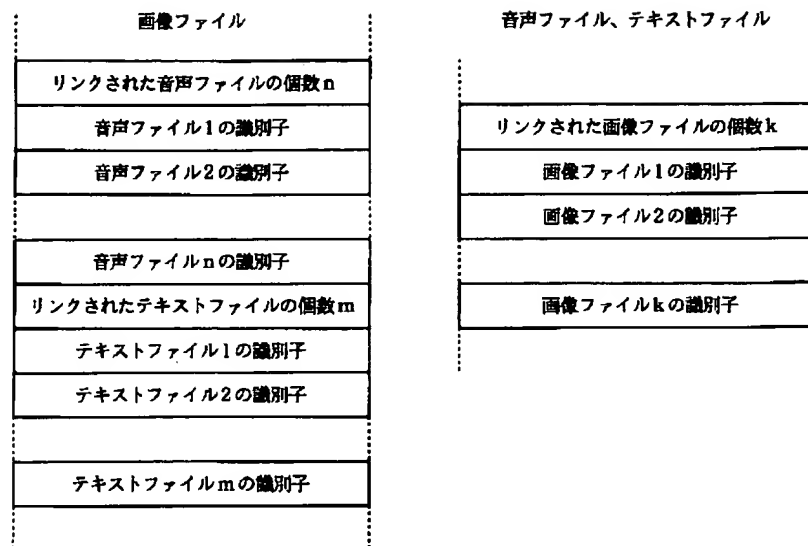
ファイル一覧

[Drawing 6]

qwertyuiopasdfghjklzxcvbnm.

閉じる —602

601

[Drawing 10]

グループファイル

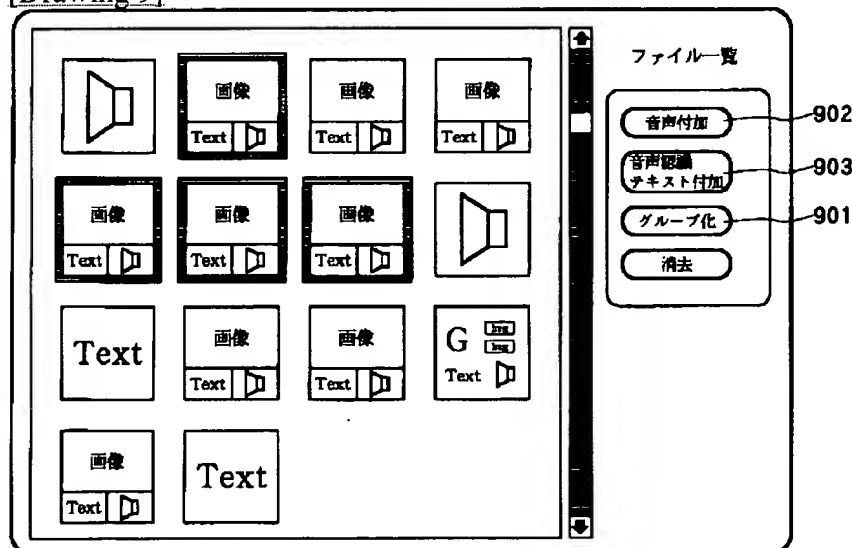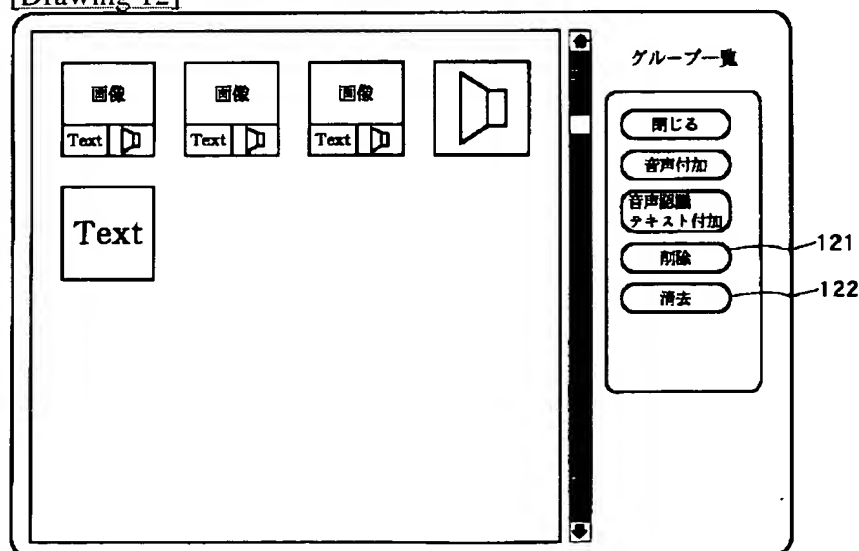| |
|---|
| リンクされた画像ファイルの個数 k |
| 画像ファイル1の識別子 |
| 画像ファイル2の識別子 |
| |
| 画像ファイル k の識別子 |
| |
| リンクされた音声ファイルの個数 n |
| 音声ファイル1の識別子 |
| 音声ファイル2の識別子 |
| |
| 音声ファイル n の識別子 |
| リンクされたテキストファイルの個数 m |
| テキストファイル1の識別子 |
| テキストファイル2の識別子 |
| |
| テキストファイル m の識別子 |

[Drawing 7]

[Drawing 8]

画像ファイル
リンクされた音声ファイルの個数 n
音声ファイル1の識別子
音声ファイル2の識別子
音声ファイル n の識別子
リンクされたテキストファイルの個数 m
テキストファイル1の識別子
テキストファイル2の識別子
テキストファイル m の識別子

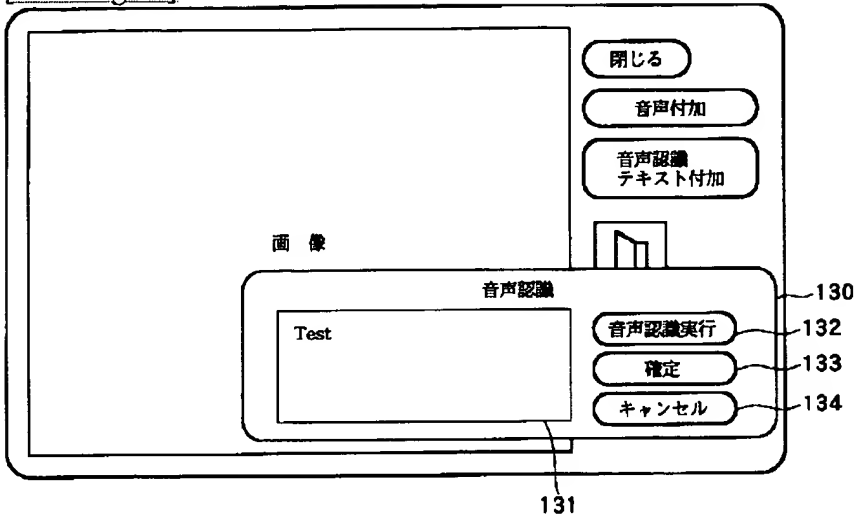| 音声ファイル、テキストファイル |
| --- |
| リンクされた画像ファイルの個数 k |
| 画像ファイル1の識別子 |
| 画像ファイル2の識別子 |
| |
| 画像ファイル k の識別子 |

[Drawing 9]

[Drawing 12]

[Drawing 13]

[Translation done.]